Is Irritable Bowel Syndrome a Diagnosis of Exclusion?: A Survey of Primary Care Providers, Gastroenterologists, and IBS Experts

Brennan M.R. Spiegel, MD, MSHS[1,2,3,4,5,6], **Mary Farid, MD**[1,2], **Eric Esrailian, MD, MPH**[2,6], **Jennifer Talley, MSPH**[1,5], and **Lin Chang, MD**[1,2,5,6]

[1] Department of Gastroenterology, VA Greater Los Angeles Healthcare System, Los Angeles, California, USA

[2] Division of Digestive Diseases, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[3] Department of Health Services, UCLA School of Public Health, Los Angeles, California, USA

[4] CURE Digestive Diseases Research Center, Los Angeles, California, USA

[5] Center for Neurobiology of Stress, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

[6] UCLA/VA Center for Outcomes Research and Education, Los Angeles, California, USA

Abstract

OBJECTIVES—Guidelines emphasize that irritable bowel syndrome (IBS) is not a diagnosis of exclusion and encourage clinicians to make a positive diagnosis using the Rome criteria alone. Yet many clinicians are concerned about overlooking alternative diagnoses. We measured beliefs about whether IBS is a diagnosis of exclusion, and measured testing proclivity between IBS experts and community providers.

METHODS—We developed a survey to measure decision-making in two standardized patients with Rome III-positive IBS, including IBS with diarrhea (D-IBS) and IBS with constipation (C-IBS). The survey elicited provider knowledge and beliefs about IBS, including testing proclivity and beliefs regarding IBS as a diagnosis of exclusion. We surveyed nurse practitioners, primary care physicians, community gastroenterologists, and IBS experts.

RESULTS—Experts were less likely than nonexperts to endorse IBS as a diagnosis of exclusion (8 vs. 72%; $P < 0.0001$). In the D-IBS vignette, experts were more likely to make a positive diagnosis of IBS (67 vs. 38%; $P < 0.001$), to perform fewer tests (2.0 vs. 4.1; $P < 0.01$), and to expend less money on testing (US\$297 vs. \$658; $P < 0.01$). Providers who believed IBS is a diagnosis of exclusion ordered 1.6 more tests and consumed \$364 more than others ($P < 0.0001$). Experts only rated celiac

sprue screening and complete blood count as appropriate in D-IBS; nonexperts rated most tests as appropriate. Parallel results were found in the C-IBS vignette.

CONCLUSIONS—Most community providers believe IBS is a diagnosis of exclusion; this belief is associated with increased resource use. Experts comply more closely with guidelines to diagnose IBS with minimal testing. This disconnect suggests that better implementation of guidelines is warranted to minimize variation and improve cost-effectiveness of care.

INTRODUCTION

Irritable bowel syndrome (IBS) is a chronic disorder of gastrointestinal function that is characterized by recurrent abdominal pain and altered bowel habits in the absence of detectable organic disease (1). Although IBS is extremely prevalent, affecting up to 15% of the general adult population (2), diagnosing IBS is not always straightforward. Properly diagnosing IBS can be challenging and uncertain for several reasons, including: (i) there is currently no consistent biological marker of IBS, leaving clinicians to rely on patient symptoms alone to make the diagnosis; (ii) the symptoms of IBS are often difficult to quantify objectively; (iii) symptoms can vary among individuals with IBS, and (iv) many organic conditions can masquerade as IBS (3). The last fact is the most troubling to clinicians and patients, many of whom remain unsettled by the prospect of overlooking alternative diagnoses such as inflammatory bowel disease, microscopic colitis, infectious colitis, small intestinal bacterial overgrowth, celiac sprue, gluten sensitivity, and colon neoplasia, among many others. This uncertainty often prompts a series of diagnostic tests to exclude alternative etiologies. In other words, many clinicians approach IBS as a diagnosis of exclusion.

Despite the tendency of clinicians to order diagnostic tests in the face of IBS symptoms, the Rome III criteria encourage clinicians to make a positive diagnosis on the basis of symptom criteria derived from the previously validated Manning and Kruis criteria, and emphasize that IBS is not a diagnosis of exclusion despite its broad differential diagnosis (1). This recommendation is based on evidence that diagnostic testing has a generally low yield in patients fulfilling the Rome criteria who otherwise lack alarming features (4,5). A decade ago Vanner *et al.* (6) showed that Rome I-positive symptoms have a positive predictive value of 98% for IBS. That is, using diagnostic testing strategies at the time, 98% of Rome-positive patients ended up having IBS, and did not have an underlying organic condition after undergoing standard evaluations.

The previous Rome II guidelines stated that IBS can be diagnosed so long as "there are no structural or biochemical abnormalities," whereas the current Rome III guidelines maintain that IBS can be diagnosed in the absence of "alarm features," and is "often properly diagnosed without testing" (1). But these clinical asterisks are understandably difficult for many providers to reconcile, and some argue that they raise more questions than they answer. Questions like: What diagnostic tests, if any, should be performed before clinicians can reliably diagnose IBS? Is it really possible to make the diagnosis without any information from diagnostic testing? Or should a complete blood count, chemistry panel, and stool occult blood test, at the very least, be drawn in all potential IBS patients? What about testing for celiac sprue, gluten sensitivity, or colon cancer?

To better understand current diagnostic decision-making in IBS, we performed a vignette survey to measure provider beliefs about whether IBS is a diagnosis of exclusion, and to measure beliefs about the appropriateness of commonly available diagnostic tests in IBS. We hypothesized that providers who believe IBS is a diagnosis of exclusion conduct more diagnostic tests and expend more resources in evaluating patients with IBS symptoms compared with providers who do not share this belief. We further hypothesized that provider

beliefs regarding diagnostic testing in IBS exhibit extreme variation beyond chance alone and diverge from published guidelines.

METHODS

Vignette survey design

We developed an online vignette-based questionnaire to evaluate decision-making in two standardized patients with bowel symptoms that met Rome III criteria for IBS: one patient met criteria for IBS with diarrhea (D-IBS), the other for IBS with constipation (C-IBS). We developed the vignettes in concert with IBS experts and survey design specialists to ensure clinical face validity, comprehensibility, and comprehensiveness — a process we have previously followed in other vignette surveys (7–10). The vignettes began with a patient history and physical examination, and were followed by management questions pertaining to diagnostic testing, treatment, and follow-up. The questions included vertical single best answers, horizontal matrix items, and open-ended items. The D-IBS vignette described a 42-year-old woman with longstanding frequent, loose stools, lower abdominal cramping that improved with stool passage, and no evidence of alarming signs or symptoms. The C-IBS vignette described a 52-year-old woman with longstanding intermittent lower abdominal pain and constipation with bloating, hard stools, straining, and a sense of incomplete evacuation after defecation. Both patients met Rome III criteria for IBS, lacked alarm features, were not receiving medications known to cause IBS symptoms, and had no family history of gastrointestinal malignancy or inflammatory bowel disease. The full vignettes are presented in the Appendix A. The vignettes were accompanied by a series of stand-alone questions pertaining to IBS etiology and clinical decision-making, including beliefs about the etiology of IBS, belief about whether IBS is a diagnosis of exclusion, and knowledge about the yield of common diagnostic tests, as described further, below. Respondents received $10 for completing the survey. The study was approved by the Veteran Administration Greater Los Angeles Institutional Review Board and was conducted in accordance with the institutional guidelines regulating human subject research.

Sampling frame

We surveyed four provider groups:

Sample of IBS key opinion leaders ("experts")—We surveyed 50 international gastroenterologists (GIs) who are recognized experts in IBS. We identified these experts based on publication records over the past 10 years, membership in practice guideline committees including the Rome Foundation Working Groups, and participation in advisory councils for the American College of Gastroenterology and the American Gastroenterological Association. We use the term "expert" in reference to this group throughout the remainder of this paper.

Simple random sample of community gastroenterologists—We surveyed a random sample of 300 GIs from the membership directory of the American Gastroenterological Association. In case the random selection process identified an IBS expert already included in the first group of providers, we repeated the selection process to identify a second individual to avoid duplicates between samples. We used a random number generator to identify subjects in this and other groups, below.

Simple random sample of general internal medicine physicians—We surveyed a random sample of 300 general internal medicine physicians (GIMs), including internists and family practitioners, from the membership directory of the American Medical Association.

Simple random sample of nurse practitioners—To include a control group of nonphysicians who regularly manage IBS, we sought a group of nurse practitioners (NPs) working in primary care. We surveyed a random sample of 300 NPs from the National Veteran Affairs provider database.

Survey sample size, distribution, and follow-up procedures

Assuming 15 subjects for each of 10 potential independent predictors in multivariable regression analysis (see "Analyses" section), we required a minimum of 150 subjects to complete the survey to avoid overmatching of the regression models. Respondents initially received the survey electronically using an online questionnaire platform (Survey Monkey software, www.surveymonkey.com). Physicians received e-mails with cover letters and a link to the online survey. After 2 weeks, nonresponders received a follow-up e-mail. Finally, 1 week after the second e-mail correspondence, a paper version of the questionnaire was mailed to nonresponders.

Analyses

Measuring variations in diagnostic decision-making—We measured the level of agreement among provider groups for endorsement of commonly available diagnostic tests for D-IBS symptoms. Following the patient history and physical examination, and in the absence of additional diagnostic testing information, respondents were asked the following question: "Based upon the information provided to this point, do you believe that this patient has irritable bowel syndrome? " Respondents selected between three options: "Yes," "No," and "Unsure —need more information," Those who believed the patient has IBS were asked a follow-up question to explore their willingness to inform the patient of the presumed diagnosis: " In addition to believing that this patient has IBS, are you also prepared at this time to confidently and affirmatively inform her that she has IBS? " Respondents selected between three options: "Yes—I am prepared to make a confident and affirmative diagnosis and inform the patient; " "No—I need more information before making a confident and affirmative diagnosis;" or "No —I am confident of the diagnosis, but I am not yet prepared to inform the patient."

Respondents were then asked to rate the appropriateness of a range of laboratory, radiographic, and endoscopic tests using a standard nine-point RAND/UCLA Appropriateness Scale (RAS) with the following interpretation: scores 1–3="generally inappropriate;" 4–6="neither inappropriate nor appropriate;" 7–9="generally appropriate" (11). We compared mean RAS scores among provider groups using analysis of variance and a *P* value <0.05 as significant. To quantify the level of agreement, we also calculated the RAND/UCLA "Disagreement Index" (DI) for each factor (11). The DI is based on the distribution and symmetry of the scores across the nine-point RAS, and has been externally validated as a measure of variation in provider beliefs and previously used in gastroenterology provider surveys (7–9). A higher DI indicates wider spread across the nine-point scale, and lower values indicate increasing consensus. If the DI exceeds 1.0, then the distribution meets criteria for "extreme variation" in ratings—more than what would be expected by chance alone. If the DI is ≤1.0, then there is not extreme variation. The DI is calculated using a standard published equation (11).

Measuring diagnostic costs—We measured the diagnostic costs of care endorsed by each respondent in their work-up of the standardized patients. To measure this outcome, we first itemized the total number of diagnostic tests endorsed by each respondent, and then assigned a cost to each diagnostic test based on a third-party payer perspective. We obtained costs for physician services and endoscopic, radiographic, and surgical procedures from the 2008 American Medical Association Current Procedural Terminology codebook and the 2008 Medicare Fee Schedule. We then derived aggregate costs per respondent by tabulating the number and type of tests performed, and then summing the assigned costs for the tests. We

compared aggregate costs expended between provider groups using analysis of variance, and also measured differences between expert and nonexpert groups using a *t*-test. We then performed multivariate regression analysis to determine if any provider or practice-type characteristics (e.g., provider age, gender, practice setting, years of practice, provider type, percent of time dedicated to clinical care, research experience, number of patients with IBS evaluated per month) were associated with overall diagnostic testing expenditures.

Belief about IBS as a diagnosis of exclusion—We posed a stand-alone question separate from the clinical vignette, as follows: "Based on your clinical experience, do you believe that IBS is primarily a diagnosis of exclusion (i.e. one or more diagnostic tests should be performed before diagnosing IBS)?" Respondents selected between two options: (i) "Yes —IBS is primarily a diagnosis of exclusion; and," or (ii) "No—IBS is not primarily a diagnosis of exclusion." We first calculated the proportion of respondents in each provider group who endorsed IBS as a diagnosis of exclusion, and then compared proportions among groups using a χ^2-test. We then performed multivariable regression analysis to evaluate the impact of this belief on diagnostic testing outcomes, including the total number of tests endorsed and the total expenses incurred in diagnostic testing. We hypothesized that respondents who endorsed IBS as a diagnosis of exclusion would order more tests and incur higher diagnostic costs than respondents who did not endorse this belief, independent of other factors.

Knowledge about yield of common diagnostic tests in IBS—Data indicate that the yield of common diagnostic tests in IBS, including colonoscopy, stool studies, inflammatory markers, and thyroid function testing, among others, is low (5,12). It is possible that some providers overestimate the diagnostic yield of common diagnostic tests in IBS, and that knowledge about test yield may provide diagnostic decision-making. We therefore posed a series of questions to assess knowledge and experience with diagnostic testing yield in IBS. Respondents viewed a list of common diagnostic tests beneath the following instructions: " Based on your clinical experience with patients similar to this woman, what percentage of the time do each of the following tests identify a treatable cause for the bowel symptoms (please provide your best estimate)." For each diagnostic test, respondents selected between three estimates of diagnostic yield: (1) < 1% of the time; (2) 1–10% of the time; and (3) > 10% of the time. In each case current published data indicate yield rates of 10% or less (5), suggesting that the optimal evidence-based response for each item was either 1 or 2, but not 3. We operationalized provider diagnostic testing knowledge with a " knowledge index," which was a composite score representing the proportion of nine diagnostic test yield questions answered correctly by each respondent —a technique we have used in previous vignette surveys (7,9). We compared mean scores among provider groups, and evaluated the bivariate and multivariate relationship between testing knowledge and diagnostic resource use outcomes. We hypothesized that less knowledge about test yield would be associated with higher diagnostic resource use.

RESULTS

Sample characteristics

Table 1 shows the characteristics of the survey respondents. A total of 308 respondents returned their surveys, including 27 of 50 IBS experts (54% response rate), 90 of 300 community GIs (30% response), 89 of 300 GIMs (30 % response), and 102 of 300 NPs (34 % response). Using data from the American Medical Association Masterfile, we found no difference between responder and nonresponders for age, gender, or years in practice. The expert group was significantly more likely to be engaged in conducting research and evaluated more IBS patients per month. However, compared with the nonexpert groups, experts had a smaller proportion of time dedicated to clinical care.

Diagnostic decision-making in D-IBS

Figure 1 shows the data regarding willingness to diagnose IBS in the D-IBS vignette. Seventy percent of the IBS experts believed the patient had IBS based on review of the history and physical examination alone, of whom 53% were prepared to confidently and affirmatively inform the patient that they have IBS. The remaining IBS experts requested additional information, described further below. Among community GIs, GIMs, and NPs, the percentages who believed the patient had IBS were 52, 34, and 41%, respectively ($P < 0.01$). Overall, IBS experts were significantly more likely to believe the patient had IBS vs. non-experts (70 vs. 39 %; $P < 0.002$). Among the 39% of nonexperts who believed the patient had IBS, only 18% were prepared to inform the patient of this diagnosis without additional diagnostic testing.

Figure 2 provides the appropriateness ratings for common diagnostic tests available in the D-IBS vignette. IBS experts only rated complete blood count (mean RAS=7.4) and serologic test for celiac sprue (RAS=6.8) as "generally appropriate," and rated lactulose breath testing for small intestinal bacterial overgrowth as "generally inappropriate" (RAS=2.3). All other tests were scored, on average, as "neither appropriate nor inappropriate." Community GI ratings closely mirrored expert ratings except for stool leukocyte measurement, which was deemed "generally appropriate" (RAS=6.0) by the community GIs. The mean scores for GIM and NP groups were higher than GI providers for most tests.

Table 2 provides the mean RAS, DIs, and P values for t-tests comparing mean scores between IBS expert vs. combined non-expert groups. Experts were consistent in their disfavor of breath testing (DI=0.5) and support of complete blood count (DI=−1.4) and celiac sprue testing (DI=−3.1), but exhibited extreme variation (i.e., DI > 1.0) for all other tests. Nonexperts revealed a mix of extreme and acceptable variation across tests. Compared with nonexperts, IBS expert ratings were significantly higher for celiac sprue testing (P=0.0002), and significantly lower for most other studies. In the D-IBS vignette, experts performed fewer diagnostic tests (2 vs. 4.1; P<0.01) and to expended less cost on testing ($297 vs. $658; P<0.01).

Diagnostic decision-making in C-IBS

Figure 3 shows the data regarding willingness to diagnose IBS in the C-IBS vignette. A total of 67% of experts believed the patient had IBS, of whom 44 % were prepared to inform the patient that he/she has IBS. Among community GIs, GIMs, and NPs, the percentages who believed the patient had IBS were 48, 17, and 12%, respectively (P<0.0001). IBS experts were more likely to believe the patient had IBS compared to nonexperts (72 vs. 20 %; P<0.0001). Among the 20% of nonexperts who believed the patient had IBS, only 24% were prepared to inform the patient of this diagnosis without additional diagnostic testing.

Figure 4 provides the appropriateness ratings for common diagnostic tests available in the C-IBS vignette. All provider groups rated complete blood count and colonoscopy as highly appropriate in this 52-year-old patient. With the exception of the expert group, all groups rated thyroid function testing as highly appropriate (P=0.0001). All other tests were scored, on average, as "neither appropriate nor inappropriate."

Table 2 provides the mean RAS, DI, and P values for t-tests for the C-IBS vignette. Experts revealed extreme variation in opinion regarding the appropriateness of anorectal manometry (DI=1.6), serum chemistry panel (DI = 1.9), and thyroid function testing (DI = 6.9). Compared with nonexperts, experts were less likely to endorse most other studies. In the C-IBS vignette, experts performed fewer diagnostic tests (1.4 vs. 2.2; P=0.05) and expended less cost on testing ($298 vs. $550; P<0.03).

Belief about IBS as a diagnosis of exclusion

Irritable bowel syndrome experts were less likely than non-experts to endorse IBS as a diagnosis of exclusion (8 vs. 72%; $P < 0.0001$). Among nonexperts, 42% of GIs and 76% of both GIMs and NPs shared this belief ($P < 0.01$ for GI vs GIM + NP). In regression analysis adjusting for provider group, expert status, age, and gender, providers who believed IBS is a diagnosis of exclusion ordered 1.6 more tests and consumed $364 more than those who did not ($P < 0.0001$) in the D-IBS vignette. In the C-IBS vignette this belief was independently associated with 1.3 more tests and $212 more in diagnostic expenditures ($P = 0.03$).

Knowledge about yield of common diagnostic tests in IBS

Table 3 provides the results regarding knowledge of diagnostic test yield in IBS. Overall, the percentage of diagnostic test knowledge items answered correctly was 88, 78, 70, and 38% for the expert, community GI, GIM, and NP groups, respectively (analysis of variance, $P<0.0001$). Across groups, respondents were most likely to overestimate the yield of a complete blood count. In multivariable regression providers with lower "knowledge index " scores used more diagnostic tests ($P=0.001$) and expended more money in diagnostic evaluations ($P=0.005$).

DISCUSSION

Diagnostic testing remains rampant in IBS despite the dissemination of the Rome guidelines. To better understand current diagnostic decision-making in IBS, we performed a survey to measure provider beliefs about whether IBS is a diagnosis of exclusion, and, if so, what diagnostic tests to use.

Our survey has six key findings: first, we found that only 8 % of experts endorsed IBS as a diagnosis of exclusion, whereas 72 % of community providers shared this belief (7). This indicates a considerable disconnect between academic guidelines and community practice. Of note, our survey cannot definitively answer whether or how this disconnect should be remedied. On the one hand, the disconnect may reflect poor dissemination and implementation of evidence-based guidelines, such as those developed by the American College of Gastroenterology (12). This would suggest that better implementation of guidelines is warranted to minimize variation and improve cost-effectiveness of care. On the other hand, the disconnect may reflect valid concerns among community physicians that our knowledge remains highly incomplete, and that underlying conditions may persist. Indeed, data indicate that gluten sensitivity (13) and microscopic colitis (14), in particular, may be more prevalent in D-IBS than previously thought, and that treatment of these entities might resolve IBS symptoms. And although commonly available diagnostic tests for small intestinal bacterial overgrowth remain inaccurate (15,16), there are groups of IBS patients whose symptoms may be precipitated or maintained due to abnormal host – microbiome interactions. In other words, despite more than a decade of guidelines indicating that IBS should be a diagnosis of exclusion, cumulating data indicate that it may yet be premature to render this judgment. Our survey results may reflect the current vulnerable state of uncertainty despite dissemination of well-reasoned, evidence-based guidelines.

Second, we found that providers who believe IBS is a diagnosis of exclusion ordered more diagnostic tests and consumed between $200 and $400 more in diagnostic expenditures per patient compared with those who did not share this belief; this finding was consistent in both D-IBS and C-IBS patients. Of note, this relationship remained significant after adjusting for a range of provider-level characteristics and was consistent across provider groups.

Third, using a standardized definition of "appropriateness," we found that experts only rated celiac sprue screening and a complete blood count as generally appropriate first-line routine

tests in D-IBS, whereas in a 52-year-old C-IBS patient only complete blood count and colonoscopy were considered appropriate by experts. In contrast, community providers rated complete blood count, chemistry panel, stool leukocytes, and stool ova and parasites as appropriate in D-IBS, whereas in C-IBS these providers rated serum chemistry panel, colonoscopy, complete blood count, and thyroid-stimulating hormone level as appropriate. Because primary care providers are generally first to evaluate IBS patients in the community, these findings may have significant economic implications since they suggest that front-line providers are especially prone to order a wide variety of tests with low diagnostic yield.

Fourth, we found that many providers overestimate the yield of common diagnostic test in IBS. This propensity is especially evident for complete blood count, thyroid function testing, and various stool studies. Despite extensive data that these tests generally do not alter the diagnosis of IBS, it appears that many community providers either do not apply these findings toward their diagnostic decisions or are not aware of them. Moreover, we have shown that knowledge of diagnostic yield independently predicts diagnostic expenditures. Future educational programs and interventions should emphasize the evidence regarding diagnostic yield; this may lower costs and streamline diagnostic decision-making in IBS.

Fifth, we revealed that experts harbored extreme variation in their opinion about the appropriateness of many tests, including erythrocyte sedimentation rate, thyroid function testing, stool studies, and anorectal manometry. The experts were even varied in their opinion of colonoscopy in a D-IBS patient under 50 years of age. In short, experts were more likely than nonexperts to comply with published diagnostic testing guidelines, but both groups showed variations from guidelines and internal inconsistencies in their beliefs about diagnostic testing.

Finally, we found that among providers willing to proactively diagnose IBS on the basis of history and physical examination alone, less than half are willing to inform the patient of their belief until additional testing has nonetheless been performed. This may reflect uncertainty about the diagnosis, medical–legal concerns, or both; our survey did not distinguish. Future research should study this phenomenon further as it may provide insight not only for management of IBS, but possibly for other chronic illnesses with wide differential diagnoses; namely, we should determine why providers may be willing to secure an "internal diagnosis" without "externalizing" the diagnosis to the patient before potentially unnecessary diagnostic testing.

Why do providers continue to order tests in IBS, despite data that these tests are generally low yield? In light of the medical– legal interface in the United States, one possibility is that some clinicians believe that diagnostic testing is a form of inoculation against litigation. But clearly this is a suboptimal reason to pursue diagnostic testing for any reason, and data from our group and others indicate that the quality of the physician–patient relationship is a critical predictor of outcomes (17), and a more important predictor of litigation than testing proclivity (18). A second possibility is the belief that even negative diagnostic tests are useful, because they can allay patient concerns about serious illness and provide reassurance. Yet we have shown that a negative colonoscopy, in particular, is not associated with reassurance or improved quality of life in young IBS patients (19). In fact, we found a nonsignificant trend toward less reassurance in patients receiving a negative colonoscopy (vs. no colonoscopy at all). Nonetheless, every patient is different, and some have fears and concerns that simply will not cease until one or more diagnostic tests return as "normal." A third possibility is that IBS patients with high levels of somatization —a process marked by multiple unexplained somatic complaints and physical illnesses potentially related to underlying psychosocial distress—are sometimes misclassified as having underlying organic conditions, and subsequently undergo sequential diagnostic tests to chase the symptoms (20–22). We previously found a linear and

highly significant relationship between levels of somatization and the amount of diagnostic testing in IBS (22), suggesting that providers should remain on the lookout for somatization in IBS, and aggressively treat or refer somatization patients in lieu of performing potentially unnecessary diagnostic tests. But perhaps the most common reason for diagnostic testing in IBS is that the Rome criteria have a 98%—not 100%—positive predictive value (6). So no matter how strong the evidence that diagnostic testing is low yield, there is always the real (although small) possibility of underlying organic disease. With evolving data and testing strategies, the positive predictive value of Rome criteria may indeed fall below the decade-old estimate from Vanner *et al.* The fact that some IBS patient have another explanation for their symptoms is simply not debatable, particularly in light of evolving data that IBS patients are a heterogeneous population with a core of " pure IBS " surrounded by small subsets of alternative diagnoses such as celiac sprue, gluten sensitivity, bacterial overgrowth, and other masqueraders (4).

Yet despite this reality, clinicians should keep in mind that time is on their side in IBS. In the absence of alarming features, the IBS masqueraders are typically chronic conditions with slowly evolving natural histories. Moreover, cancer is no more common in IBS patient than controls (5), and patients more than 50 years of age should receive colorectal cancer screening regardless of IBS symptoms. Assuming there are no alarming signs or symptoms, clinicians might benefit from focusing less on diagnostic testing and more on addressing disease-specific fears and concerns, setting mutually reasonable goals and expectations, providing multimedia educational materials (such as those offered by the American Gastroenterological Association or the International Foundation for Functional GI Disorders), teaching self-empowerment techniques, empirically treating symptoms, and screening for and treat ing (or refer) somatization, as suggested by the American College of Gastroenterology IBS guidelines (12). But ultimately, clinicians must use their judgment, and should reserve the right to investigate further if their IBS patients do not " follow the script," either because of a poor response to therapy, worsening symptoms over time, unabated patient concerns that undermine quality of life, or development of incident alarming features (3). Like most things in medicine, diagnostic testing in IBS remains a balance of art and science.

A potential limitation of this study is that survey responses may not be reflective of actual decision-making in clinical practice. Directly observing patient –provider interactions is considered the gold standard for assessing process of care. However, this approach is also limited because of the Hawthorne effect in which providers artificially alter their practice when they are knowingly observed. This undermines the efforts to capture the true process of care. Standardized patients (23) and medical record data abstraction (24) are alternatives. Notably, survey-based clinical vignettes have been validated as an accurate surrogate for both chart abstraction and standardized patients (25), and are thus widely recognized to be a valid, reliable, practical, and cost-effective technique to assess process of care. An additional limitation is that we did not track geographic location or method of payment for professional services of the respondents. These factors might affect diagnostic decision-making, as both are associated with resource utilization in other areas of medicine (26), as we have found in dialysis (27). In addition, we did not measure respondent knowledge about the Rome criteria—it is possible that this knowledge might influence clinical decision-making. A further limitation is that although our C-IBS vignette is consistent with Rome criteria, it could also represent painful constipation, an entity described by Drossman *et al.* (28) and Bharucha *et al.* (29). Whereas C-IBS implies a temporal association between pain and constipation, "painful constipation" indicated a coexistence of the symptoms, but not necessarily a temporal relationship. It is unclear, however, whether this distinction would alter providers' diagnostic testing strategies in a meaningful way. Finally, our distinction of expert vs. nonexpert, although based on explicit criteria, may fail to acknowledge the fact that many community providers who manage IBS on a daily basis might have more clinical experience than academic thought leaders. Moreover,

our findings do not necessarily mean that community providers are providing suboptimal, poor, or inadequate care—only that their decisions are at time discordant with published evidence and guidelines. It is indeed possible that published evidence and guidelines are themselves shortsighted, faulty, or incomplete; only time will tell.

Our survey is further limited by the relatively low response rate. There may be systematic differences between responders and non-responders. Nonresponders typically cite a wide range of reasons for their failure to participate, including not receiving the survey, believing the survey did not pertain to their practice, lacking sufficient time, or simply choosing not to participate. We cannot know the specific reasons for nonresponse to this survey. It is possible that nonresponders were too busy, did not receive the surveys, or had other systematic differences—i.e., do not evaluate enough patients to feel comfortable taking the survey, do not participate in academic studies, are not clinically active, etc. Using data from the American Medical Association Masterfile, we have confirmed that non-responders were similar to responders in terms of age, gender, and years in practice. This suggests that respondents were not systematically different from nonrespondents for key measured characteristics. Nonetheless, there may be unmeasured factors, including those described above, that remain systematically different between groups. As with any provider survey, a higher response rate could have potentially altered the results, although it is difficult to predict accurately in what direction, if any, the results would be different. We have expanded our discussion to include these comments.

Overall, we found that best practice diagnostic guideline have not been uniformly adopted in IBS, particularly among primary care providers; persistent guideline – practice disconnects should be addressed. High rates of diagnostic resource use may be driven by potentially modifiable knowledge deficits, beliefs, and attitudes regarding IBS. Specifically, resource use is higher among providers who believe IBS is a diagnosis of exclusion, and among those who have less complete knowledge about the yield of common diagnostic tests. Given the high population prevalence and extraordinary costs attributed to IBS, these findings have important implications for how to guide care in a more cost-effective direction. Investigators should develop and implement health system interventions and educational approaches to improve adherence with current evidence-based best practices in IBS diagnostic testing by focusing on the specific areas identified in this survey. We should also remain cognizant that future data may refine current diagnostic guidelines, and that diagnostic testing strategies must evolve as new data emerge.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- Current guidelines emphasize that irritable bowel syndrome (IBS) is not a diagnosis of exclusion and encourage clinicians to make a positive diagnosis using the Rome criteria alone.

- Yet many clinicians are concerned about overlooking alternative diagnoses.

WHAT IS NEW HERE

- Best practice diagnostic guidelines have not been uniformly adopted in IBS, particularly among primary care providers.

- Most community providers believe IBS is a diagnosis of exclusion; this belief is associated with increased diagnostic resource use.

- Despite the dissemination of guidelines regarding diagnostic testing in IBS, there remains extreme variation in beliefs among both experts and nonexperts.

> - This disconnect suggests that better implementation of guidelines is warranted to minimize variation and improve cost-effectiveness of care.

Acknowledgments

References

1. Longstreth GF, Thompson WG, Chey WD, et al. Functional bowel disorders. Gastroenterology 2006;130:1480–91. [PubMed: 16678561]

2. Spiegel BM. The burden of IBS: looking at metrics. Curr Gastroenterol Rep 2009;11:265–9. [PubMed: 19615301]

3. Spiegel B. Do physicians follow evidence-based guidelines in the diagnostic work-up of IBS? Nat Clin Pract Gastroenterol Hepatol 2007;4:296–7. [PubMed: 17541444]

4. Cash BD, Schoenfeld P, Chey WD. The utility of diagnostic tests in irritable bowel syndrome patients: a systematic review. Am J Gastroenterol 2002;97:2812–9. [PubMed: 12425553]

5. Cash BD, Schoenfeld P, Chey WD. The utility of diagnostic tests in irritable bowel syndrome patients: a systematic review. Am J Gastroenterol 2002;97:2812–9. [PubMed: 12425553]

6. Vanner SJ, Depew WT, Paterson WG, et al. Predictive value of the Rome criteria for diagnosing the irritable bowel syndrome. Am J Gastroenterol 1999;94:2912–7. [PubMed: 10520844]

7. Esrailian E, Gralnek IM, Jensen D, et al. Evaluating the process of care in nonvariceal upper gastrointestinal haemorrhage: a survey of expert vs. non-expert gastroenterologists. Aliment Pharmacol Ther 2008;28:1199–208. [PubMed: 18729846]

8. Esrailian E, Spiegel BM, Targownik LE, et al. Differences in the management of Crohn's disease among experts and community providers, based on a national survey of sample case vignettes. Aliment Pharmacol Ther 2007;26:1005–18. [PubMed: 17877507]

9. Spiegel BM, Farid M, van Oijen MG, et al. Adherence to best practice guidelines in dyspepsia: a survey comparing dyspepsia experts, community gastroenterologists and primary-care providers. Aliment Pharmacol Ther 2009;29:871–81. [PubMed: 19183152]

10. Spiegel BM, Ho W, Esrailian E, et al. Controversies in ulcerative colitis: a survey comparing decision making of experts versus community gastroenterologists. Clin Gastroenterol Hepatol 2009;7:168–74. 174, e1. [PubMed: 18952199]

11. Fitch, K.; Bernstein, SAM., et al. The RAND/UCLA Appropriateness Method User's Manual. RAND; Santa Monica: 2001.

12. Brandt LJ, Chey WD, Foxx-Orenstein AE, et al. An evidence-based position statement on the management of irritable bowel syndrome. Am J Gastroenterol 2009;104 (Suppl 1):S1–35. [PubMed: 19521341]

13. Wahnschaffe U, Schulzke JD, Zeitz M, et al. Predictors of clinical response to gluten-free diet in patients diagnosed with diarrhea-predominant irritable bowel syndrome. Clin Gastroenterol Hepatol 2007;5:844–50. quiz 769. [PubMed: 17553753]

14. Limsui D, Pardi DS, Camilleri M, et al. Symptomatic overlap between irritable bowel syndrome and microscopic colitis. Inflamm Bowel Dis 2007;13:175–81. [PubMed: 17206699]

15. Corazza GR, Menozzi MG, Strocchi A, et al. The diagnosis of small bowel bacterial overgrowth. Reliability of jejunal culture and inadequacy of breath hydrogen testing. Gastroenterology 1990;98:302–9. [PubMed: 2295385]

16. Gasbarrini A, Corazza GR, Gasbarrini G, et al. Methodology and indications of H2-breath testing in gastrointestinal diseases: the Rome Consensus Conference. Aliment Pharmacol Ther 2009;29 (Suppl 1):1–49. [PubMed: 19344474]

17. Spiegel BMRNB, Mayer E, Bolus R, et al. The effectiveness of a model physician–patient relationship versus usual care in irritable bowel syndrome: a randomized controlled trial. Gastroenterology 2006:A773.

18. Vincent C, Young M, Phillips A. Why do people sue doctors? A study of patients and relatives taking legal action. Lancet 1994;343:1609–13. [PubMed: 7911925]

19. Spiegel BM, Gralnek IM, Bolus R, et al. Is a negative colonoscopy associated with reassurance or improved health-related quality of life in irritable bowel syndrome? Gastrointest Endosc 2005;62:892–9. [PubMed: 16301033]

20. Miller AR, North CS, Clouse RE, et al. The association of irritable bowel syndrome and somatization disorder. Ann Clin Psychiatry 2001;13:25–30. [PubMed: 11465682]

21. Whitehead WE, Palsson O, Jones KR. Systematic review of the comorbidity of irritable bowel syndrome with other disorders: what are the causes and implications? Gastroenterology 2002;122:1140–56. [PubMed: 11910364]

22. Spiegel BM, Kanwal F, Naliboff B, et al. The impact of somatization on the use of gastrointestinal health-care resources in patients with irritable bowel syndrome. Am J Gastroenterol 2005;100:2262–73. [PubMed: 16181379]

23. Colliver JA, Swartz MH. Assessing clinical performance with standardized patients. JAMA 1997;278:790–1. [PubMed: 9286845]

24. Luck J, Peabody JW, Dresselhaus TR, et al. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. Am J Med 2000;108:642–9. [PubMed: 10856412]

25. Peabody JW, Luck J, Glassman P, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. Ann Intern Med 2004;141:771–80. [PubMed: 15545677]

26. The Dartmouth Atlas of Healthcare. Feb. 2010 Accessed at: http://www.dartmouthatlas.org/

27. Desai AA, Bolus R, Nissenson A, et al. Is there "cherry picking" in the ESRD Program? Perceptions from a Dialysis Provider Survey. Clin J Am Soc Nephrol 2009;4:772–7. [PubMed: 19339407]

28. Drossman DA, Morris C, Hu Y, et al. Further characterization of painful constipation (PC): clinical features over one year and comparison with IBS. J Clin Gastroenterol 2008;42:1080–8. [PubMed: 18949838]

29. Bharucha AE, Locke GR, Zinsmeister AR, et al. Differences between painless and painful constipation among community women. Am J Gastroenterol 2006;101:604–12. [PubMed: 16464225]

APPENDIX A

Clinical vignette no. 1 (D-IBS)

Patient history

A 42-year-old white woman visits your office with a longstanding history of intermittent abdominal pain and diarrhea that has become progressively worse over the previous 4 months. She describes the pain as "crampy" and usually in the left lower quadrant of the abdomen. The pain often improves when she passes a bowel movement and is worse within 10 min of eating a meal. She passes 4–8 bowel movements a day that are usually "loose" and "come on urgently." She describes his appetite as "good," but mentions that she sometimes avoids eating because it brings on her symptoms. She denies weight loss. She has no history of fevers, chills, sweats, vomiting, or rectal bleeding. She has no recent travel history, unusual food ingestions, or antibiotic use. She denies intolerance to dairy products. Her medical history is significant only for an appendectomy at age 22. She has no allergies and takes no medications. She reports social alcohol use. Her father has type II diabetes mellitus, and her 32-year-old sister has chronic diarrhea of undiagnosed origin.

Physical examination

She is 5′8″ and 170 lbs (77 kg). Her vital signs are normal. She has mild tenderness to palpation in the left lower quadrant of the abdomen. Her stool is guaiac negative. The remainder of her examination is normal.

No additional information is elicited at this time. In light of the information provided, please answer the questions that follow to the best of your ability, even if you would normally elicit more information.

Clinical vignette no. 2 (C-IBS)

Patient history

A 52-year-old woman visits your office with a longstanding history of intermittent abdominal pain and constipation that has become progressively worse over the previous 8 months. She describes the pain as "dull and achy," usually accompanied by bloating, and located in the lower abdomen. She passes only two bowel movements per week that are usually "hard." She complains that she must strain for "a couple of minutes" before spontaneously evacuating her stool, although manual disimpaction has not been necessary. She also complains of a feeling of "incomplete emptying" after defecation. Her stool caliber has not changed. She has no history of fevers, chills, sweats, vomiting, weight loss or rectal bleeding. She was prescribed a fiber supplement by her primary physician, but complained that it caused bloating and did not improve her symptoms. She has a history two normal spontaneous vaginal deliveries without complications. She has no allergies and takes no medications other than ibuprofen as needed for knee pain. She reports social alcohol use. There is no family history of gastrointestinal cancer or other significant medical conditions.

Physical examination

She is 5′6″ and 125 lbs (57 kg). Her vital signs are normal. She has mild tenderness to palpation in the center of the abdomen. Her rectal examination reveals firm stool in the vault that is guaiac negative. Her sphincter tone and perineal descent are normal. The remainder of her examination is normal.

No additional information is elicited at this time. In light of the information provided, please answer the questions that follow to the best of your ability, even if you would normally elicit more information.
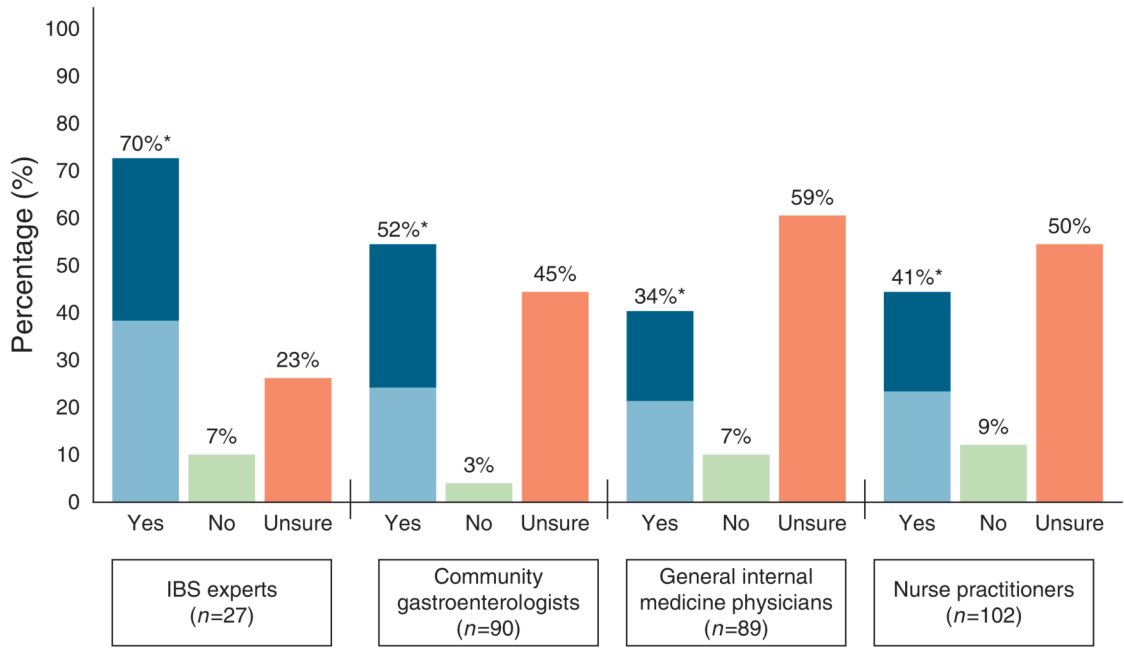
Figure 1.
Willingness to diagnose irritable bowel syndrome (IBS) on the basis of patient history and physical examination fi ndings alone in D-IBS vignette. Providers were asked: "Based upon the information provided to this point, do you believe that this patient has irritable bowel syndrome?" The fi gure depicts the percentages by group responding "Yes," "No," or "Unsure —need more information." Among those answering "Yes," we further explored willingness to inform the patient of the presumed diagnosis by asking: "In addition to believing that this patient has IBS, are you also prepared at this time to confi dently and affi rmatively inform her that she has IBS?" The light blue of the "Yes" bar demonstrates the subset prepared to make a confi dent and affi rmative diagnosis without further diagnostic testing.
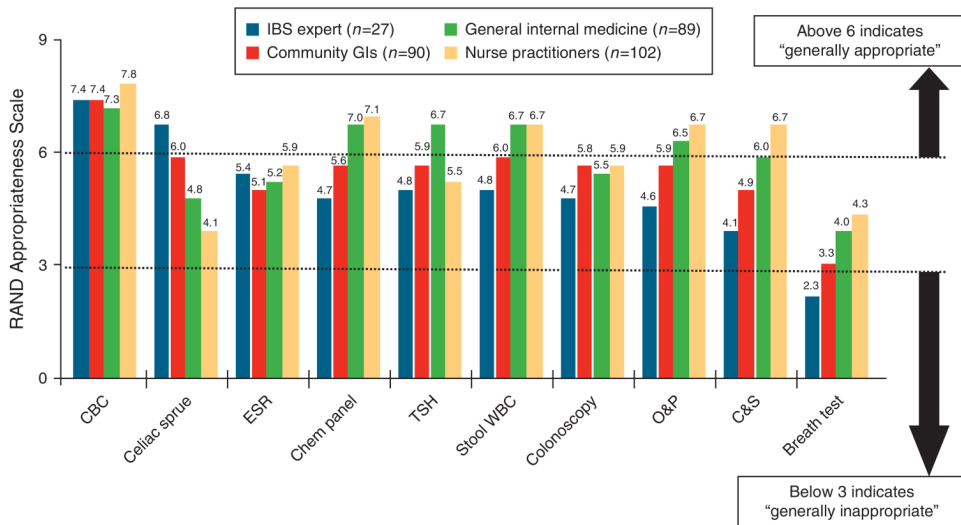
Figure 2.
Provider appropriateness ratings for performing common diagnostic tests in D-IBS vignette. Respondents scored the appropriateness of each test using a standard nine-point RAND/UCLA Appropriateness Scale, in which a score of 1–3 is "generally inappropriate," 4–6 is "neither appropriate nor inappropriate," and 7–9 is "generally appropriate." Mean scores exceeding 6 indicate that providers believe, on average, that a test is generally appropriate; mean scores below 3 indicate that providers believe a test is generally inappropriate. CBC, complete blood count; ESR, erythrocyte sedimentation rate; D-IBS, irritable bowel syndrome with diarrhea; O&P, ova and parasites; TSH, thyroid-stimulating hormone; WBC, white blood cells.
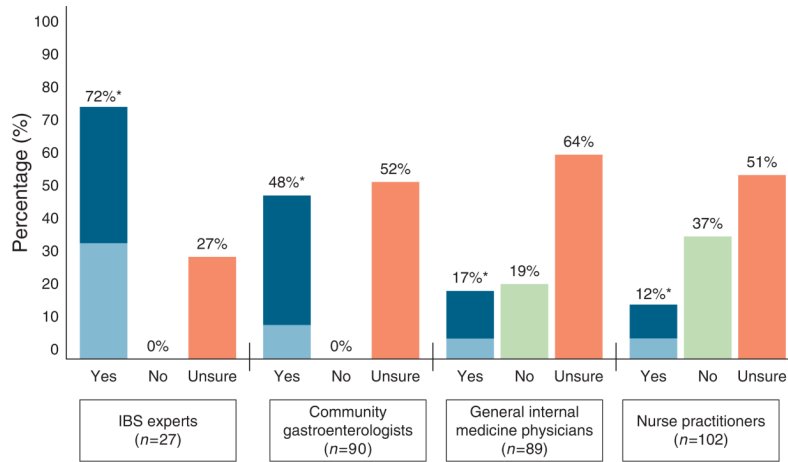
Figure 3.
Willingness to diagnose irritable bowel syndrome (IBS) on the basis of patient history and physical examination fi ndings alone in C-IBS vignette. IBS experts were more likely to believe the patient had IBS vs. nonexperts (72 vs. 20%; $P < 0.0001$). Refer to Figure 1 legend for interpretation. *White bar indicates percentage of experts who believe the patient has IBS. Light blue bar indicates sub-percentage who both believes patient has IBS and is willing to inform the patients without further testing.
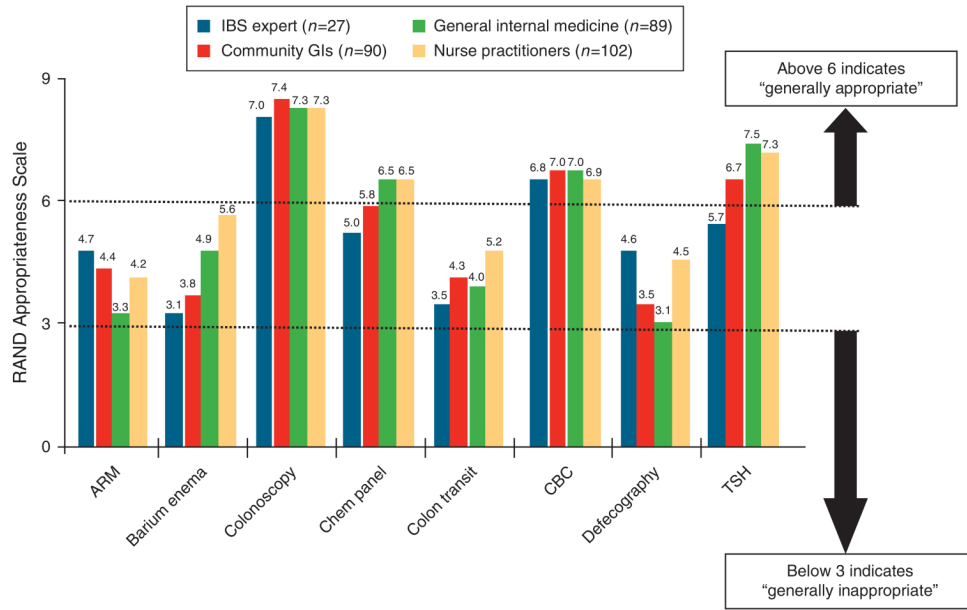
Figure 4.
Provider appropriateness ratings for performing common diagnostic tests in C-IBS vignette. Refer to Figure 2 legend for interpretation. ARM, anorectal manometry; CBC, complete blood count; TSH, thyroid-stimulating hormone.

Table 1

Demographic and practice-pattern information of respondents

Variable	IBS experts (*n*=27)	Community GIs (*n*=90)	Primary care providers (*n*=89)	Nurse practitioners (*n*=102)	*P* value
Age (mean years ± s.d.)	49±10	47±9	47±10	50±8	0.23
Male gender (%)	74	81	57	4	<0.001
Years in practice	26±11	25±9	26±15	21±12	0.01
Time dedicated to clinical practice (%)	36	79	83	92	<0.001
Time dedicated to research (%)	39	7	5	1	<0.001
Respondents who evaluate/treat > 20 IBS patients per month (%)	41	17	0	1	<0.001

GI, gastroenterologist; IBS, irritable bowel syndrome.

Table 2

Comparison of mean RAS scores and DIs between IBS expert vs. nonexpert provider groups for common D-IBS and C-IBS diagnostic tests

Test	IBS experts		IBS nonexperts		
	Mean RAS	DI	Mean RAS	DI	*P* value
D-IBS vignette					
Breath test for bacterial overgrowth	2.3	0.5	4.1	**4.1**	0.0001
Celiac sprue serologies	6.8	−3.1	4.8	**1.7**	0.0002
Chemistry panel	4.7	**1.1**	6.9	−6.2	< 0.0001
Colonoscopy	4.7	**1.6**	5.9	**7.0**	0.04
Complete blood count	7.4	−1.4	7.5	−0.9	0.7
Erythrocyte sedimentation rate	5.4	**4.1**	5.4	**1.8**	0.94
Stool culture and sensitivity	4.1	**1.6**	6.2	**1.6**	0.0001
Stool leukocytes	4.8	**1.7**	6.6	−6.2	0.0003
Stool ova and parasites	4.6	**1.6**	6.6	−6.2	0.0001
Thyroid-stimulating hormone level	4.8	**1.8**	6.2	−12	0.006
C-IBS vignette					
Anorectal manometry	4.7	**1.6**	3.9	0.7	0.08
Barium enema	3.1	0.5	2.3	**1.7**	0.0004
Chemistry panel	5.0	**1.9**	6.5	**30**	0.004
Colonoscopy	7.0	−0.9	7.4	−0.9	0.29
Colon transit study (e.g., radio-opaque marker study)	3.5	0.7	4.6	0.9	0.03
Complete blood count	6.8	−2.8	6.9	−2.1	0.75
Defecography study	3.0	0.4	3.7	0.7	0.11

Test	IBS experts		IBS nonexperts		
	Mean RAS	DI	Mean RAS	DI	P value
Thyroid-stimulating hormone level	5.7	**6.9**	7.3	−7.4	0.0001

C-IBS, IBS with constipation; DI, disagreement index; D-IBS, IBS with diarrhea; IBS, irritable bowel syndrome; RAS, RAND/UCLA Appropriateness Scale.

A DI ≥1.0 indicates "extreme variation" in beliefs within a group (bold values), whereas a DI < 1.0 indicates increasing consensus. *P* values compare mean RAS scores between groups.

Table 3

Knowledge of diagnostic yield of common tests for IBS symptoms

Diagnostic test	IBS experts, n=27 (%)	Community GIs, n=90 (%)	Primary care providers, n=89 (%)	Nurse practitioners, n=102 (%)	P value
Overall percentage correct (of 9 diagnostic yield knowledge questions)	88	78	70	38	< 0.0001
Erythrocyte sedimentation rate	93	84	80	46	< 0.001
Serology for celiac sprue	93	81	68	33	< 0.001
Colonoscopy	89	61	50	16	< 0.0001
Stool culture and sensitivity	89	85	73	36	< 0.001
Stool leukocytes	89	73	68	39	< 0.001
Serum chemistry panel	88	81	77	50	< 0.001
Thyroid-stimulating hormone	88	81	80	48	< 0.001
Stool assay for *Clostridium difficile*	85	84	63	28	< 0.0001
Complete blood count	81	70	73	40	< 0.001

GI, gastroenterologist; IBS, irritable bowel syndrome.

Data reveal the proportion of subjects from each group who correctly responded that diagnostic yield of each test in IBS is < 10% for identifying a treatable cause for the bowel symptoms. See text for further details.